

PAT-NO: JP410255058A  
DOCUMENT-IDENTIFIER: JP 10255058 A  
TITLE: MOVEMENT DETECTOR AND MOVEMENT DETECTING  
METHOD  
PUBN-DATE: September 25, 1998

INVENTOR-INFORMATION:  
NAME  
NAKAI, HIROAKI

ASSIGNEE-INFORMATION:  
NAME COUNTRY  
TOSHIBA CORP N/A

APPL-NO: JP09061392

APPL-DATE: March 14, 1997

INT-CL (IPC): G06T007/20, G06T007/00

ABSTRACT:

PROBLEM TO BE SOLVED: To provide a movement detector which detects a specific movement of a person, etc., by using an image pickup means of a TV camera, etc.

SOLUTION: A movement detector 10 comprises an image accumulating part 11 which accumulates plural images as time series images, a motion vector calculating part 12 which calculates a motion vector of the time series images, a probability distribution calculating part 13 which calculates the probability distribution of calculated motion vectors and a movement detecting part 14 which detects a specific movement from the motion vector and the probability distribution that is preliminarily calculated automatically detects

specific

movement by fetching motion information from time series images of person

movement, etc., and performing statistical learning and also greatly reduces

error detection except the movements of an object to be detected.

COPYRIGHT: (C)1998, JPO

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-255058

(43) 公開日 平成10年(1998) 9月25日

(51) Int.Cl.<sup>6</sup>

G 0 6 T 7/20

7/00

識別記号

F I

G 0 6 F 15/70

4 1 0

4 6 5 A

審査請求 未請求 請求項の数 6 O L (全 6 頁)

(21) 出願番号

特願平9-61392

(22) 出願日

平成9年(1997) 3月14日

(71) 出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72) 発明者

中井 宏章

大阪府大阪市北区大淀中1丁目1番30号

株式会社東芝関西支社内

(74) 代理人

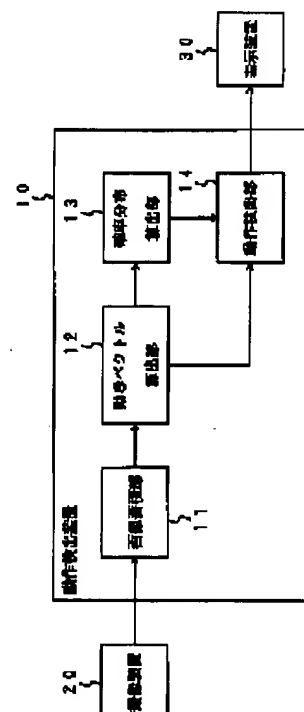
弁理士 鈴江 武彦 (外6名)

(54) 【発明の名称】 動作検出装置および動作検出方法

(57) 【要約】

【課題】 TVカメラ等の撮像手段を用いて人物等の特定の動作を検出する動作検出装置装置を提供する。

【解決手段】 動作検出装置 10 は、複数の画像を時系列画像として蓄積する画像蓄積部 11 と、時系列画像での動きベクトルを算出する動きベクトル算出部 12 と、算出された動きベクトルの確率分布を算出する確率分布算出部 13 と、動きベクトルと予め求められた確率分布とから特定の動作を検出する動作検出部 14 とからなり、人物動作等の時系列画像から動き情報を取り出して統計学習することによって、特定の動作を自動検出できるようにするとともに検出対象の動作以外の誤検出を大幅に減少させる。



1

## 【特許請求の範囲】

【請求項1】 時系列に撮像された複数の画像から被写体の特定の動作を検出する動作検出装置であって、前記複数の画像を時系列に蓄積する画像蓄積手段と、この画像蓄積手段に蓄積された連続する画像間での予め定められた領域内の動きベクトルを算出する動きベクトル算出手段と、

前記特定の動作に関する動きベクトルの確率分布を記憶する確率分布記憶手段と、

前記動きベクトル算出手段で算出された動きベクトルと前記確率分布手段で記憶された確率分布とから得られる評価値によって前記特定の動作を検出する動作検出手段とを具備してなることを特徴とする動作検出装置。

【請求項2】 時系列に撮像された複数の画像から被写体の特定の動作を検出する動作検出装置であって、前記複数の画像を時系列に蓄積する画像蓄積手段と、前記画像内で監視対象とする被写体を検出し、この検出結果に基づいて動作検出対象領域を設定する動作検出対象領域設定手段と、

前記画像蓄積手段に蓄積された連続する画像間での前記動作検出対象領域内の動きベクトルを算出する動きベクトル算出手段と、

前記特定の動作に関する動きベクトルの確率分布を記憶する確率分布記憶手段と、

前記動きベクトル算出手段で算出された動きベクトルと前記確率分布手段で記憶された確率分布とから得られる評価値によって前記特定の動作を検出する動作検出手段とを具備してなることを特徴とする動作検出装置。

【請求項3】 前記動きベクトル算出手段で算出された動きベクトルからその確率分布を算出する確率分布算出手段をさらに具備してなることを特徴とする請求項1または2記載の動作検出装置。

【請求項4】 前記確率分布記憶手段は、前記特定の動作以外の他の動作に関する動きベクトルの確率分布を前記特定の動作に関連づけて記憶する手段を具備し、前記動作検出手段は、前記他の動作に関する評価値を用いて前記特定の動作の評価値を下方に適性化する手段を具備してなることを特徴とする請求項1、2または3記載の動作検出装置。

【請求項5】 時系列に撮像された複数の画像から被写体の特定の動作を検出する動作検出方法であって、前記特定の動作に関する動きベクトルの確率分布を取得するステップと、

前記複数の画像を時系列に蓄積するステップと、前記時系列に蓄積された連続する画像間での予め定められた領域内の動きベクトルを算出するステップと、

前記算出された動きベクトルと前記取得された確率分布とから得られる評価値によって前記特定の動作を検出するステップとからなることを特徴とする動作検出方法。

【請求項6】 時系列に撮像された複数の画像から被写

2

体の特定の動作を検出する動作検出方法であって、前記特定の動作に関する動きベクトルの確率分布を取得するステップと、

前記複数の画像を時系列に蓄積するステップと、

前記画像内で監視対象とする被写体を検出し、この検出結果に基づいて動作検出対象領域を設定するステップと、

前記時系列に蓄積された連続する画像間での前記動作検出対象領域内の動きベクトルを算出するステップと、

前記算出された動きベクトルと前記取得された確率分布とから得られる評価値によって前記特定の動作を検出するステップとからなることを特徴とする動作検出方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、たとえばTVカメラ等の撮像手段を用いて人物等の特定の動作を検出する動作検出装置に関する。

【0002】

【従来の技術】TVカメラ等の撮像手段を用いて人物の手振り動作等のジェスチャを自動検出・認識する画像処理技術は、コンピュータやロボット、電化製品等への動作指示を効率よく行う目的や、挙動不審者の監視や介護が必要な病人等の映像モニタリングの目的で適用することができ、その実現が強く望まれている。

【0003】また、これまでに実験室内等の制限された撮影条件の下でのジェスチャ認識等の技術が開発されているが、これらは一般の室内環境（事務室や家庭、工場、病院等）や屋外のような環境変動の大きな場合には、認識対象となるジェスチャ以外のノイズによる誤認識が多く、実用には適さなかった。

【0004】

【発明が解決しようとする課題】このように、従来の撮像手段を用いた人物動作の自動検出・認識技術では、認識対象となる動作以外のノイズによる誤検出が多いため、一般環境での実用が困難であった。

【0005】本発明はこのような事情を考慮してなされたものであり、撮像手段により得られた人物動作等の時系列の画像から動き情報を取り出して統計学習することによって、特定の動作を自動検出できるようにするとともに検出対象の動作以外の誤検出を大幅に減少させることを可能とする動作検出装置および動作検出方法を提供することを目的とする。

【0006】

【課題を解決するための手段】本発明の動作検出装置は、複数の画像を時系列に蓄積する画像蓄積手段と、この画像蓄積手段に蓄積された連続する画像間での動きベクトルを算出する動きベクトル算出手段と、前記特定の動作に関する動きベクトルの確率分布を記憶する確率分布記憶手段と、前記動きベクトル算出手段で算出された動きベクトルと前記確率分布手段で記憶された確率分布

とから得られる評価値により前記特定の動作を検出する動作検出手段とからなるものである。

【0007】本発明によれば、撮像手段により得られた人物動作等の時系列の画像から動き情報を取り出して統計学習することによって、特定の動作を自動検出できるようにするとともに検出対象の動作以外の誤検出を大幅に減少させることが可能となる。

【0008】

【発明の実施の形態】以下、図面を参照して本発明の実施形態を説明する。

（第1実施形態）まず、この発明の第1実施形態を説明する。図1は、第1実施形態に係る動作検出装置の概略構成を示す図である。図1に示すように、この第1実施形態の動作検出装置10は、画像蓄積部11と、画像蓄積部11に蓄積されている時系列画像（時系列に撮像された複数の画像）での動きベクトルを算出する動きベクトル算出部12と、動きベクトル算出部12にて算出された動きベクトルからその確率分布を算出する確率分布部13と、動きベクトル算出部12で求められる動きベクトルと予め確率分布算出部13にて求められた確率分布とから動作を検出する動作検出部14とからなる。この動作検出装置10では、撮像装置20で逐次撮像される画像を時系列画像として画像蓄積部11に蓄積する。また、動作検出部14が検出対象とする特定の動作を検出したときは、その旨を表示装置30により表示する。

【0009】そして、この動作検出装置10は、検出すべき動作を予め学習するモード（学習モード）と、学習結果を用いて逐次撮像される映像から動作を検出するモード（検出モード）との2つのモードをもつ。

【0010】まず、学習モードについて説明する。この処理の流れを図2に示す。最初に、画像蓄積部11に時系列画像が蓄積される（ステップA1）。ここでは、画像蓄積部11に貯えられている時系列画像を  $I(x, y, t)$  と表す（図3参照）。 $x, y$  は画像中の画素の \*  $F(x, y, t) = (u(x, y, t), v(x, y, t))$

【0014】学習モード時の確率分布算出部13では、まず動きベクトル算出部12で求められた全ての動きベクトル  $F(x, y, t)$  の  $x$  軸および  $y$  軸方向成分のヒストグラム  $h(u, v)$  を求める。動きベクトルの総数を  $N$ （時系列画像  $I(x, y, t)$  の全画素について動きベクトル  $F(x, y, t)$  を求めた場合  $N = X \times Y \times T$  ※

$$P_a(u, v) = \frac{1}{N} h(u, v) = \frac{1}{N} \sum_{x=1}^X \sum_{y=1}^Y \sum_{t=1}^T \langle F(x, y, t) = (u, v) \rangle$$

【式2】

$$\text{ここで } \langle m = n \rangle = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}$$

【0016】このようにして、いま学習する対象の動作  $a$  に関する確率分布  $P_a(u, v)$  が求められれば（ステップA3）、確率分布算出部13でこの  $P_a(u, v)$  \*

\*座標  $(x: [1 \cdots X], y: [1 \cdots Y])$  であり、 $t$  はその画像が撮影された時刻あるいはフレーム番号  $(t: [1 \cdots T])$  である。この時系列画像は、撮像装置20で撮影される画像の全画面であってもよいし、予め定められた領域であってもよい。以下の学習モードの説明では、映像中で人がある動作  $a$ （例えば図3のように手を振っている等）を行っている画像領域に時系列画像  $I(x, y, t)$  が設定されているものとする。

【0011】動きベクトル算出部12で求められる動きベクトル  $F(x, y, t)$  は、時系列画像の2つの連続するフレーム  $I(x, y, t)$  と  $I(x, y, t)$  間での画像の動き、言い換えれば時刻  $t$  での座標  $(x, y)$  の画素での画像の動きを表す（図4参照）。この動きベクトルは、別名オプティカルフローとも呼ばれるが、Horn & Schunkの方法に代表される濃淡画像の時空間勾配から求める方法と、MPEG等の映像圧縮に用いられるような相関演算によるブロックマッチングによって求める方法がある（参考文献：D. H. Ballard, C. M. Brown 著 福島他訳 コンピュータサイエンス研究書シリーズ29「コンピュータ・ビジョン」日本コンピュータ協会）。

【0012】動きベクトル算出部12では、任意の動きベクトル算出方法のうちの1つを用いて画像蓄積部11に貯えられている時系列画像  $I(x, y, t)$  の動きベクトル  $F(x, y, t)$  を算出する（ステップA2）。この時、動きベクトルは  $I(x, y, t)$  の全画素について求めてもよいし、予め定められた画素のみについて求めてもよい。動きベクトル  $F(x, y, t)$  の  $x$  軸方向成分および  $y$  軸方向成分を各々  $u(x, y, t)$  および  $v(x, y, t)$  とすれば、動きベクトル  $F(x, y, t)$  は以下の式【式1】で表される2次元ベクトルである。

【0013】

【数1】

【式1】

※  $T$ ）とした場合、学習モード時に人が行っている動作  $a$  についての動きベクトルの確率分布  $P_a(u, v)$  は以下の式【式2】で求められる。

【0015】

【数2】

★  $v$ ）を記憶しておき、学習モードは終了する。学習された確率分布の例を図5に示す。この図では、手を左右に振っている人の動作を学習したものを想定し、図中  $a$  の

部分が確率値の高い部分を示している。

【0017】以上では、ある1人の動作のみについて学習する流れを説明したが、その人が動作aを何度も行ったり、複数の人が動作するのを画像蓄積部11に蓄積しておき、それらをまとめて時系列画像I(x, y, t)として学習モードを行うことにより、より多くの動作aの例について学習することができる。

【0018】次に検出モードでの処理の流れを説明する。この処理の流れを図6に示す。画像蓄積部11には、学習モード時と同様に、撮像装置20からの映像が 10 画像列I'(x, y, t)(x: [1...X'], y: [1...Y'], t: [1...T'])として蓄積されている(ステップB1)。

【0019】ただし、撮像装置20において撮影された映像の全画面中から動作を検出するために、映像中の任

$$Q_a = H_a(P_a(F')) = H_a(P_a(u', v')) \quad [式3]$$

【0022】例えば、もっとも単純に全ての動きベクトルについての動作aらしさの確率Paの和として評価関数Haを定めるならば、Haは以下のような式[式4]

$$H_a(p) = \sum_{\forall F'} p = \sum_{x=1}^{X'} \sum_{y=1}^{Y'} \sum_{t=1}^{T'} p \quad [式4]$$

また、動きベクトルが0ベクトルのときを無視するような確率Paの結合確率として評価値Qaを求めるのであれば、Haは以下のような式[式5]となる。

$$H_a(p) = \prod_{\forall F' : F' \neq (0, 0)} p \quad [式5]$$

【0025】このようにして、評価関数Haによって求められた動作aらしさの評価値Qaが、予め定められたしきい値thをこえたときに(ステップB4のY)、時系列画像I'(x, y, t)は動作aの画像列であるとして(ステップB5)、動作aの検出モードを終了する。このとき、撮像装置20からの映像において設定する全ての場所に時系列画像I'(x, y, t)を設定し、評価値Qaがしきい値thを超える場所を全て求めて、動作aを全画面中で探索してもよいし、全てのI'(x, y, t)の評価値Qaのうちで最大であって、かつ、しきい値thを超えるもの1つを検出してもよい。

【0026】(第2実施形態)次にこの発明の第2実施形態を説明する。この第2実施形態では、ある動作aのみにて学習するのではなく、他の複数の動作あるいは検出対象外の画像の動きb, c...を含めて学習し、動作検出の精度を高める方法について述べる。

【0027】学習モードは、第1の実施形態と同様であるが、確率分布算出部13では動作aの確率分布Pa(u, v)だけを求めるのではなく、他の複数の動作

$$Q' a = H' a(Qa, Qb, Qc, \dots) \quad [式6]$$

【0030】ここでの評価関数H'aは、検出目的の動作aの評価値Qaが大きな値のときにはQ'aが大きくなり、検出目的外の他のQbやQcが大きいときには

\*意の場所に画像列I'(x, y, t)が設定される。動きベクトル算出部12でも、学習モード時と同様に画像列I'(x, y, t)における動きベクトルF'(x, y, t)が求められ(ステップB2)、また、確率分布算出部13では予め学習モード時に学習された動作aの確率分布Pa(u, v)が記憶されている。

【0020】動作検出部14では、学習モード時の確率分布Pa(u, v)と、検出モード時の動きベクトルF'(x, y, t) = (u'(x, y, t), v'(x, y, t))とから、時系列画像I'(x, y, t)の動作aらしさを表す評価値Qaを、Haを任意の評価関数として以下の式[式2]で求める(ステップB3)。

【0021】

【数3】

※になる。

【0023】

【数4】

★【0024】

【数5】

☆b, c...の確率分布Pb(u, v), Pc(u, v), ...も同時に学習する。他の動作についての学習は、第1実施形態での動作aの学習と全く同様に、画像蓄積部11に別々に蓄積されている動作a, b, c, ...の時系列画像について、動きベクトル算出部12で動きベクトルを算出し、確率分布算出部13で各々の確率分布Pa(u, v), Pb(u, v), Pc(u, v), ...を別々に求め、学習モードを終了する。

【0028】検出モードでは、各動作a, b, c, ...の評価値Qa, Qb, Qc, ...が任意の評価関数Ha, Hb, Hc, ...により第1実施形態と全く同様に求められるものとする。この第2実施形態では、これらの評価値を組み合わせて最終的な評価値Q'を求めることによって検出の精度を高める。ここで、任意の関数H'aを用い、以下の式[式6]で動作aに関する最終評価値をQ'aを求める。

【0029】

【数6】

◆Q'aが小さくなるような関数であることが特徴である。ここでのb, cの学習のように、その効果が評価値を低める方向に働く学習を「負の学習」と呼ぶ。動作a

が他の動作や画像の動き  $b, c, \dots$  と紛らわしい場合でも、このように負の学習を行うことにより、精度よく動作  $a$  のみを検出することができる。学習動作が  $a, b, c$  の3つの時の  $H'$   $a$  の最も単純な例としては、以下の\*

$$H' a (Qa, Qb, Qc) = Qa - Qb - Qc$$

【0032】なお、前述した実施形態では、映像中の任意の場所に少なくとも一つ以上の画像列  $I' (x, y, t)$  を予め設定しておく例を示したが、たとえば映像中の人物を検出する手段 (TVカメラなどが撮像した画像から検出するものであってもよいし、画像を撮像するTVカメラなどと併設されるセンサなどによって検出するものであってもよい) を備え、その検出結果に基づいて動的に画像列  $I' (x, y, t)$  を設定することも有効である。

【0033】

【発明の効果】以上詳述したように、本発明の動作検出装置および動作検出方法によれば、従来の方法では検出目的以外の動作や画像の動きによる誤検出が多いために目的とする動作の検出が困難な環境においても、人物動作等の時系列画像から動き情報を取り出して統計学習することによって、特定の動作を自動検出できるようになるとともに検出対象の動作以外の誤検出を大幅に減少さ※

\* ようなもの [式7] があげられる。

【0031】

【数7】

【式7】

※ せることが可能となる等の実用上多大なる効果が奏せられる。

【図面の簡単な説明】

【図1】本発明の第1および第2実施形態の動作検出装置の概略構成を示す図。

【図2】同実施形態の動作検出装置の学習モードでの処理の流れを示すフローチャート。

【図3】同実施形態の時系列画像を説明するための図。

【図4】同実施形態の動きベクトルを説明するための図。

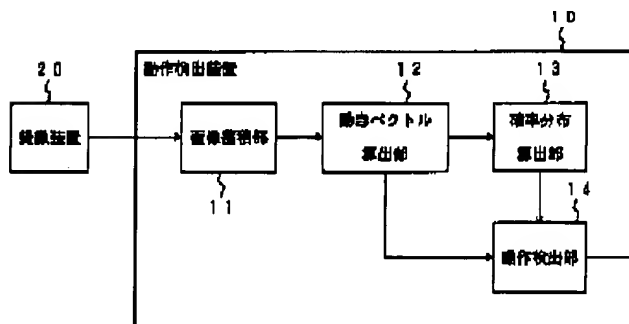
【図5】同実施形態の確率分布の一例を示す図。

【図6】同実施形態の動作検出装置の検出モードでの処理の流れを示すフローチャート。

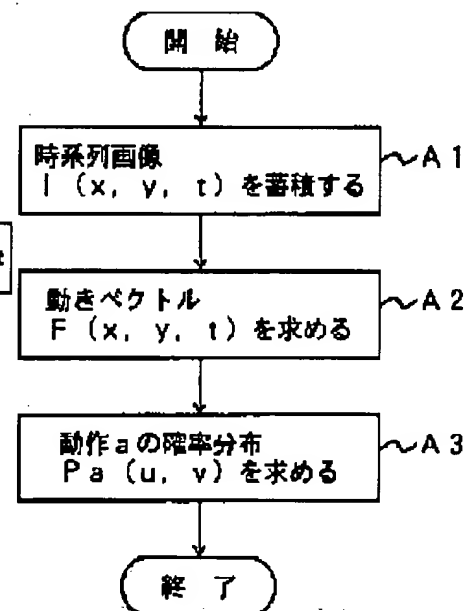
【符号の説明】

10…動作検出装置、11…画像蓄積部、12…動きベクトル算出部、13…確率分布算出部、14…動作検出部、20…撮像装置、30…表示装置。

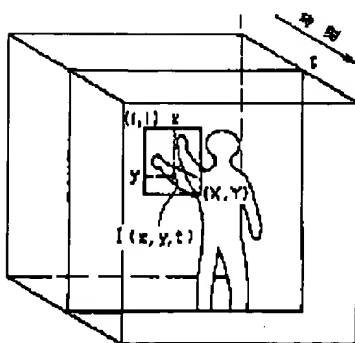
【図1】



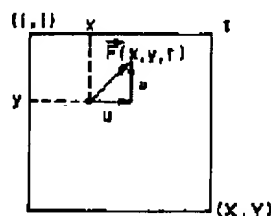
【図2】



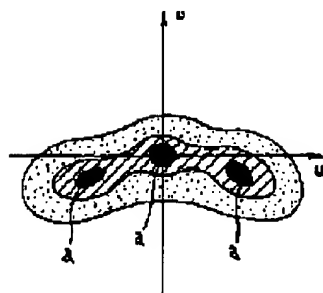
【図3】



【図4】



【図5】



【図6】

